ITEM 20 : LA MÉTHODOLOGIE DE LA RECHERCHE EN SANTÉ

RECHERCHE EN SANTÉ

Définitions

Recherche en santé

- Toute étude ou expérimentation organisée ou pratiquée sur l'être humain vivant en vue du **développement des connaissances biologiques ou médicales**
→ Développement d'actions susceptibles d'améliorer la santé
- **Pour décrire un phénomène de santé, l'expliquer, mettre au point ou évaluer un test diagnostique ou pronostique :** évaluation de l'efficacité thérapeutique des médicaments ou autres stratégies thérapeutiques, évaluation de la performance de tests diagnostiques, identification de facteurs de risque ou pronostiques, évaluation de formes d'organisation des soins ou de politiques sanitaires…
- Organisée selon trois niveaux en France :
  . **Décision** et définition de la politique de recherche (national), orientations stratégiques, organisations et allocation des moyens
  . **Programmation** pour la répartition des moyens entre grandes agences
  . Mise en œuvre de la rechercher par des **opérateurs**, unités de recherche, laboratoires (CNRS, universités, etc.)

Méthodologie

- Organisation rationnelle d'une démarche pour aboutir à un résultat
- Etapes élémentaires à mettre en œuvre pour conduire la recherche et répondre à l'objectif principal
- Ces étapes seront décrites très précisément en amont du démarrage, par écrit, dans un **protocole** de recherche, qui ne devrait **pas être modifié** en cours d'étude

Acteurs

Promoteur

= **Personne physique ou morale** prenant l'initiative de **conduire la recherche**, assurant sa **gestion** et vérifiant que le **financement** est prévu : **établissement de santé, institution** (organisme de recherche public) ou **industriel** (compagnie pharmaceutique). **Responsable légal de la recherche.**

Missions

- Obtenir l'avis **du CPP**
- Obtenir les **autorisations de l'autorité compétente = ANSM ou ABM**
- S'assurer que la recherche est **financée**
- **Sélectionner les investigateurs**
- Souscrire une **assurance responsabilité civile** spécifique (RIPH1 & RIPH2)
- Informer CPP et ANSM des **événements indésirables,** de la **fin de la recherche**

Investigateur

= **Personne physique** dirigeant et surveillant la **réalisation de la recherche sur un lieu conformément au protocole**
- **Investigateur principal** sur un lieu, auquel sont associés des **collaborateurs** de l'investigateur : médecins ou non, désignés par écrit, exerçant sous sa surveillance
- **Investigateur coordinateur** (si recherche multicentrique)

Missions

- **Inclure les sujets** se prêtant à la recherche après les avoir informés et obtenus leur consentement oral et écrit si nécessaire
- Se conformer au **guide européen des bonnes pratiques de recherche clinique**
- **Recueillir les données**
- **Administrer les traitements ou interventions** conformément au protocole
- **Déclarer les événements indésirables au promoteur**

Participant

= Toute personne volontaire (patient ou volontaire sain) peut participer à une recherche clinique
- Sous réserve d'être affiliée à un régime de Sécurité sociale (RIPH1 & RIPH2, pas RIPH3)
- Sous conditions pour : femmes enceintes, parturientes, mères qui allaitent, mineurs, majeurs sous protection juridique, personnes en soins psychiatriques sans consentement, personnes privées de liberté par décision judiciaire ou administrative
- Doit recevoir une **information préalable** à son inclusion (orale +/- résumé écrit validé par CPP), **adaptée, complète, loyale, compréhensible**. Présente :
  . Les objectifs de la recherche et son déroulement
  . Les bénéfices attendus, les contraintes et risques connus
  . Les autres traitements possibles en dehors de la recherche
  . L'avis du comité de protection des personnes
  . Les modalités du traitement des données et de leur confidentialité (CNIL)
- Donne son **consentement libre et éclairé**, par écrit (RIPH1), exprès (RIPH2) ou sa **non-opposition** orale (RIPH3)
- Peut se retirer de l'étude à tout moment sans avoir à donner de raisons et sans que sa prise en charge ultérieure en soit affectée : le patient doit en être informé
- Droit d'accès et de modification aux données recueillies la concernant

Comité de protection des personnes (CPP)

- Membres du CPP désignés par le DG-ARS, pour 3 ans
- Compétences multiples (dont en protection des données), réparties en **deux collèges** :
Premier collège : professionnels de santé + 2 méthodologistes ou biostatisticiens
Deuxième collège : personnes qualifiées en droit, éthique, sciences humaines et sociales et actions sociales, représentants des usagers

Acteurs	Comité de protection des personnes (CPP)	Missions	- Vérifier que la recherche respecte les conditions fixées par la loi : **valide la qualification de la recherche**, peut modifier la qualification indiquée par le promoteur - Examiner les **conditions de protection des participants** à la recherche : qualité, exhaustivité, intelligibilité des informations délivrées aux personnes sollicitées ; procédure de recueil du consentement, justification des recherches sur les personnes vulnérables, délai de réflexion, mesures prises pour assurer la sécurité des personnes - Examiner **la validité scientifique et éthique du projet** : pertinence scientifique, adéquation de la méthodologie aux objectifs, pertinence du niveau de preuve apporté, adéquation des moyens aux objectifs, qualification des investigateurs ➜ **Avis décisionnel et non pas seulement consultatif**

ENCADREMENT JURIDIQUE DE LA RECHERCHE

Réglementation

- **Principes d'Helsinki (1964):** principes éthiques de la recherche biomédicale. Prolonge le code de Nuremberg de 1947
- **Loi de santé publique du 9 août 2004** : cadre légal de la recherche biomédicale et de toute recherche interventionnelle (y compris en soins courants), acteurs de la recherche biomédicale et leurs obligations
- **Loi Jardé de 2012 ++**
- **Code de la Santé publique**
- Données de santé : **RGPD, CNIL**
- **Attention** : toutes les études expérimentales sont interventionnelles, mais il existe des études interventionnelles qui ne sont pas expérimentales lorsqu'il n'y a pas de modification du niveau ou de la nature du facteur d'exposition

Loi Jardé (Loi n° 2012-300 du 5 mars 2012), pour les **recherches impliquant la personne humaine** (RIPH)
- Entrent *a priori* dans le champ de la loi : les recherches organisées et pratiquées sur des personnes volontaires saines ou malades, en vue du développement des connaissances biologiques ou médicales
- Ne sont pas des RIPH au sens de la Loi Jardé : les recherches qui n'ont pas pour finalité le développement des connaissances biologiques ou médicales, la plupart des recherches portant sur des **cosmétiques**, les **enquêtes de satisfaction** liées aux soins ou non, **les expérimentations en sciences humaines et sociales dans le domaine de la santé** (sauf psychothérapies et TCC : catégorie 2), les recherches **non organisées ou non pratiquées évaluant les modalités d'exercice** des professionnels de santé ou des **pratiques d'enseignement** dans le domaine de la santé ou exploitant des **données rétrospectives**.
- Ne nécessitent pas d'autorisation de CPP, mais peuvent nécessiter des démarches auprès de la CNIL

Catégories de RIPH		**1 : recherche interventionnelle**	**2 : interventionnelle à risques et contraintes minimes**	**3 : non interventionnelle (= observationnelle)**
Recherche		- Intervention sur la personne non justifiée par sa prise en charge habituelle, non dénuée de risques	- Ajout par rapport à la pratique courante d'intervention(s) dont la réalisation ne comporte que des risques et contraintes minimes - Exclut la recherche sur les médicaments et dispositifs médicaux de catégorie 1, même utilisés de manière habituelle	- Actes et produits habituels, conformes à la pratique courante, sans procédure supplémentaire ou inhabituelle de diagnostic, traitement ou surveillance, sans aucun risque ni contrainte supplémentaire.
		Donc une étude observationnelle (cas-témoins, etc.) peut être interventionnelle légalement si un examen ou une modalité de surveillance sont ajoutés par rapport à la pratique courante !		
Autorisations et obligations	ANSM	Autorisation	Information (transmission de l'avis CPP + résumé de la recherche)	
	CPP	Avis favorable		
	Information	Information individuelle du patient		
	Consentement	- **Ecrit** Nb : consentement des 2 parents si mineur	- Ecrit ou oral	- Non-opposition : signature qu'en cas de refus du patient
	CNIL	- Autorisation de la CNIL - Ou engagement de conformité MR001 (engagement à suivre les méthodologies de référence)		- Autorisation de la CNIL - Ou engagement MR002 ou MR003, voire MR001 en cas de consentement
	Promoteur	Nécessaire systématiquement		
	Assurance	Nécessaire		Non nécessaire

PRINCIPES DE LA RECHERCHE

- Mise à l'épreuve d'hypothèses de recherche ou construction de modèles
- Première étape = formulation de la question de recherche, **claire et précise**
- Repose sur des méthodes et outils

Principes

Méthodes

- Etudes diagnostiques : évaluation de nouveaux tests de diagnostic ou de dépistage
- Etudes épidémiologiques : description de l'état de santé de la population, identification des déterminants de santé, évaluation des interventions en santé
- Méthode choisie fonction de **l'objectif poursuivi** et de la maîtrise ou non du facteur/intervention :
  . Evaluer la validité et fiabilité d'un nouveau test → étude diagnostique
  . Caractériser l'état de santé d'une population → étude descriptive
  . Identifier des facteurs de risque ou facteurs pronostiques → étude analytique
  . Evaluer l'efficacité d'un traitement ou d'une stratégie thérapeutique → étude expérimentale
  . Evaluer la performance d'une organisation de soins, d'un programme de prévention ou d'une politique sanitaire → étude expérimentale ou quasi-expérimentale
- Toutes les études n'ont pas la même valeur démonstrative en termes de causalité -> **niveaux de preuve** différents *(cf. item 323)*

Etudes diagnostiques	- Toujours observationnelles - But : établir la valeur diagnostique d'un test clinique ou paraclinique, dans une démarche de dépistage ou de diagnostic
Etudes descriptives	- Toujours observationnelles - But : décrire la distribution ou l'évolution de la fréquence des événements de santé et de leurs déterminants dans la population - Pas de conclusion causale ou inférentielle, mais possibles hypothèses explicatives
Etudes explicatives	- But : identifier/quantifier/interpréter l'association entre un facteur et un événement - Deux types selon que le facteur étudié est : . Maîtrisé : étude **expérimentale** = essai clinique contrôlé randomisé . Non maîtrisé : étude **étiologique** = analytique. Cas-témoins, cohorte, exposés-non exposés, transversale, de registre

Population

- **Population cible** = celle à laquelle on souhaite extrapoler les résultats
- **Population source** = celle sur laquelle se déroule concrètement l'étude ; représentative de la population cible
- Etudes menées sur des sous-ensembles de la population source : échantillonnage

→ Détermine la **validité externe** de l'étude

Modalités échantillonnage	Sur une base de sondage : - Recrutement **exhaustif** - **Tirage su sort :** sondage aléatoire simple, sondage stratifié (tirage au sort des individus dans chaque sous-population), sondage en grappes (groupes et non individus tirés au sort) Sans sondage disponible : - **Recrutement consécutif** (au fur et à meure de leur arrivée dans les centres d'inclusion), refus de participer et non-réponse doivent être faibles ; si possible exhaustif, ou tirage au sort systématique (1/2 ou 1/3) - **Quotas :** échantillon représentatif car structure similaire à la population source ; méthode non-aléatoire, précision des estimations inconnue - **Echantillon de convenance :** à éviter (risque de non-représentativité)
Recueil des données	- Préciser la **nature**, la **source**, les **modalités de mesure** et de **recueil** - Etude *ad hoc* : échantillon d'individus interrogés et examinés - **Sources de données d'enregistrement continu** : systèmes de surveillance **épidémiologiques** : registres de morbidité, déclaration obligatoire, réseau Sentinelles, statistiques vitales (CépiDC, INSEE) - **Bases médico-administratives** : SNDS (SNIIRAM, PMSI, CépiDC, handicap), entrepôts de données de santé (EDS), système d'information hospitalier Nouvelles sources de données : - Génome humain, exposome, phénome, données –omics (biologie moléculaire), données environnementales, données issues d'objets connectés

Outils biostatistiques

Définitions

- **Probabilités :** modélisent des phénomènes aléatoires dont issues possibles connues mais dont on ne peut prédire la valeur car leur réalisation est incertaine
- **Variable aléatoire :** fonction qui attribue un nombre à toutes les issues possibles ; discrète ou continue
- **Loi de probabilité :** décrit la probabilité associée à chaque valeur (valeur discrète) ou à tout intervalle de valeur (variable continue)

Variabilité et tests

- Critères biomédicaux sont, par définition, **variables**
- **Fluctuations d'échantillonnage :**
  . Différences entre échantillons du simple fait du hasard
  . Tests statistiques pour conclure si le résultat peut être dû au hasard : **risque α de 1ère espèce** (en général 5%)

Inférence statistique

- La description porte sur un échantillon représentatif de la population cible
- **Intérêt de l'inférence = extrapolation**
- Risque de se tromper, à causes des fluctuations d'échantillonnage, quantifiable par :
  . Tests statistiques adaptés
  . Estimation des intervalles de confiance

Hypothèses

- Démarche **hypothético-déductive** = tester des hypothèses par des expériences
- Test d'hypothèse = décision entre une **hypothèse nulle H0** (on cherche à la rejeter) et une **hypothèse alternative H1** (complémentaire à H0, que l'on cherche à accepter)

→ **Risque alpha** = probabilité de rejeter H0 alors qu'elle est vraie ; souvent fixé à 5%
→ **Risque bêta** = probabilité de ne pas rejeter H0 alors que H1 est vraie
→ **Puissance** :

- Probabilité de rejeter H0 si H1 est vraie = capacité à montrer une différence ou une association si elle existe
- Varie selon la taille de l'échantillon, le risque alpha, le test uni- ou bilatéral (ces trois paramètres sont contrôlables), la différence à détecter et la variabilité des observations

Etapes du test

- Définir H0 et H1 et fixer le **risque alpha**
- Déduire ce que l'on devrait observer si H0 était vraie, et déterminer une zone de rejet de H0
- Vérifier si ce que l'on observe est conforme ou non à ce que l'on attend sous H0

→ Si ce que l'on observe est conforme : non rejet de H0 (ce qui ne veut pas dire que H0 est vraie, car le risque beta n'est jamais nul). Si H0 correspond à une égalité, on conclura qu'il n'y a **pas de différence significative.**
→ Si ce que l'on observe est non conforme : le résultat du test est le rejet de H0. Il existe toujours un risque de rejeter H0 à tort (le risque alpha), mais dans ce cas, on est autorisé à conclure en faveur de H1. Si H0 correspond à une égalité, on conclura qu'il existe une **différence significative.**

- Calculer le **degré de signification**

Type test

Quand H0 est une égalité, construction de H1 :
- H1 est **une différence** quel que soit son sens -> test **bilatéral**
- H1 est **une différence** dans un sens unique -> test **unilatéral**

Choix du test

- Différent selon que les données sont **appariées** (plusieurs mesures chez un même individu ou des individus appariés) ou **indépendantes** (chaque mesure obtenue chez un individu différent et indépendant)
- **Test paramétrique** (ne peut être utilisé que si certaines conditions sont remplies) ou **non-paramétrique** (conditions non remplies)
- Choix du test fonction du **type de critère de jugement** (variable dépendante), du **type de facteur d'exposition** (variable explicative), de la **structure des données** (indépendantes ou appariées), de la **taille des échantillons** (> ou < 30) et de la **distribution statistique des variables** (normale ou non) :

		Variable Y (à expliquer = dépendante)			
		Binaire	Qualitative nominale à > 2 classes	Quantitative	Censurée
Variable X (explicative = indépendante)	Binaire	Chi² ou Fisher		Student ou Mann-Whitney	Logrank
	Qualitative nominale > 2 classes			ANOVA Kruskall-Wallis	
	Quantitative	Student ou Mann-Whitney	ANOVA Kruskall-Wallis		

	Données de survie	- **Délai jusqu'à la survenue d'un événement** ; 2 informations : survenue ou non de l'événement + temps de suivi du sujet (= durée de survie) - **Censure** : quand l'événement ne survient pas au cours du suivi ou perdu de vue -> durée de survie = durée pendant laquelle on a pu observer le sujet ou délai jusqu'à dernière visite du sujet **Méthode de Kaplan-Meier** : - Estimation de la survie à chaque instant T - **Courbe de survie** : démarre à 1 (aucun sujet avec l'événement), et diminue au cours du temps (« marches d'escalier ») ; traits verticaux = censures ; médiane de survie = temps au bout duquel 50% des individus de l'échantillon ont connu l'événement - Analyse univariée : comparaison de deux courbes de survie (test du log-rank) - Analyse multivariée : modèle de Cox avec estimation d'Hazard Ratio (HR)

PROTOCOLE

Plan d'analyse

Aspects fondamentaux et grandes lignes du déroulement définis ***a priori*** dans le **protocole** de l'étude ; argumentaire justifiant la question de recherche, la pertinence des hypothèses explorées, la formulation des objectifs poursuivis ainsi que la méthodologie

Justification		Formulation de l'objectif et de la question ou hypothèse de recherche reposant sur une analyse de la littérature (rapporter clairement les références) : - **Arguments épidémiologiques :** fréquence, facteurs de risque, conséquences, coût du problème de santé - **Etat actuel des connaissances :** plausibilité, analogie, études préliminaires à faible niveau de preuve ou résultats contradictoires, reproductibilité, cohérence…
Objectif		Découle directement de l'hypothèse formulée : - Si plusieurs objectifs : hiérarchiser en **principal et secondaire(s)** - L'objectif principal doit être : **unique (si possible), clair, précis, consensuel, raisonnable**
Matériel et méthode	**Méthode**	**Plan d'étude mis en œuvre**, selon l'étude : - Descriptive : étude sur une période donnée, pas de groupe comparateur, estimation de la fréquence d'un phénomène de santé - Explicative : groupe comparateur pour étude du lien entre un facteur et un phénomène de santé ; randomisation du facteur = étude expérimentale (lien de causalité), sinon = étude étiologique (suspicion de causalité)
	Matériel	**Groupe d'intérêt** : - Population source dont sera issu l'échantillon étudié ; critères d'inclusion et non-inclusion (**extrapolabilité**) ; sélection doit être la plus large possible, non liée à un niveau particulier d'exposition aux facteurs étudiés - Formes cliniques et critères diagnostiques reconnus et définis précisément ; repérer un éventuel biais de sélection - Définir si **cas incidents** (nvx cas, évite le biais de survie sélective) ou **cas prévalents** (modification possible des habitudes par rapport au début de la maladie, perte de la séquence temporelle) - Préciser le lieu de recrutement des sujets et s'interroger sur la représentativité des cas **Groupe comparateur :** - Défini par la randomisation dans les études expérimentales ; comparabilité des caractéristiques initiales par la randomisation - Dans les études étiologiques, le groupe comparateur appartient à la même population d'origine que la population étudiée - Attention aux biais, au moment de la conception du protocole : choix des personnes incluses (biais de sélection), mesure des données recueillies (biais de mesure) ; discuter des mesures pour limiter ces biais - **Procédures identiques** entre les groupes : mesures réalisées de manière identique, recueil des mêmes variables ; utilisation de l'insu dans les études expérimentales **Critères de jugement** : - Un CJP doit être défini dans l'étude (calcul du NSN) - Exprimé sous forme d'une fréquence (variable qualitative) ou d'une moyenne (variable quantitative) ; critère d'efficacité, de tolérance ou une mesure d'association - **Pertinence** (signification clinique, en rapport avec l'objectif, moment du recueil, moyens fiables de recueil), **validité (**objectivité), **disponibilité potentielle** et **fiabilité** (risque de biais de mémorisation dans les études cas-témoins)… - Parfois : critères intermédiaires (moindre pertinence clinique), critères secondaires (exploration d'hypothèses)

Plan d'analyse	Matériel et méthode	**Nombre de sujets nécessaire (NSN)**	Nombre de sujets à inclure estimé pour démontrer l'hypothèse, sur la base du critère de jugement principal, calculé *a priori* Dépend : - Etude descriptive : de **la précision souhaitée de l'estimation du paramètre** → NSN d'autant plus important que l'estimation souhaitée est précise (donc l'intervalle de confiance resserré autour de l'estimation ponctuelle) - Etude explicative : de **l'importance de l'écart sur le critère de jugement** principal à montrer, du **risque de 1ère espèce α et de la puissance statistique souhaitée** (1 – ß), de la **variabilité du critère** → NSN d'autant plus important que le risque α est faible et la puissance souhaitée élevée, la variabilité du critère de jugement importante, et le différentiel attendu minime
	Analyse statistique		- Planifiée *a priori*, repose toujours sur une description de la population étudiée : . Décrit les écarts à la population prévue à l'inclusion → **extrapolable** ? . Si étude comparative : comparaison des groupes à l'inclusion, perdus-de-vue et écarts au protocole → **facteurs de confusion** ? (ajustement nécessaire) Analyses descriptives : - Description variables qualitatives : effectifs et proportions - Description variables quantitatives : . Paramètres de position : moyenne, mode (= valeur la + fréquente), quantile (médiane, quartiles…) . Paramètres de dispersion : variance et écart-type, distance interquartile, étendue **Analyse du CJP** : par des tests appropriés - Toute autre analyse est à titre documentaire car inflation du risque α - Expression des résultats en termes probabilistes, degré de signification ou *P-value* - En cas de différence significative, à interpréter : pertinence ou intérêt clinique

APPROCHES MÉTHODOLOGIQUES

Les études en santé peuvent être classées selon leur **objectif** et la **question de recherche**, la **procédure employée**, la **chronologie du design** ou le **choix des sujets**.

Enquêtes descriptives			- Permettent de décrire **l'évolution** ou **la distribution** des évènements de santé ou leurs déterminants dans la population, et de rechercher les facteurs de risque d'une maladie - Les informations sont recueillies de façon systématique ou par enquêtes spécifiques (population exhaustive ou échantillon représentatif) - **Enquêtes transversales** (prévalence) et **enquêtes longitudinales** de cohorte (incidence, survie)
	Enquêtes transversales		- Photographie de la population à un instant T, **pas de suivi** - Recueil **simultané** des caractéristiques de sujets, de l'existence d'une maladie ou d'expositions - Estimation de **la prévalence** d'un événement ou de **la fréquence** d'une exposition, avec un IC95% (cet IC ne permet pas de conclure à la significativité)
	Enquêtes de cohorte descriptives		Cohorte = ensemble d'individus avec des caractéristiques connues et semblables, **suivis** au cours du temps : nés la même année, ayant la même activité professionnelle, une même exposition ou défini sur une base géographique (*cohorte « Framingham »)* - Sujets inclus **indemnes de l'événement** étudié au moment de l'inclusion - Le plus souvent **prospectives** (information sur l'exposition recueillie **avant** que l'événement de santé ne soit survenu chez les sujets), parfois **rétrospectives** (cohortes historiques ; information sur l'exposition recueillie **après** que l'événement de santé ne soit survenu chez les sujets) - Estimation de **la mortalité** ou de **l'incidence** (taux d'incidence ou densité d'incidence) d'un événement de santé (si populations non comparables -> **standardisation**), voire **le pronostic** d'une maladie (taux de survie, létalité, récidive). Utilisation des techniques d'analyse de survie
Enquêtes analytiques			- Toujours **observationnelles** et **comparatives** - Cherchent à identifier **l'existence d'une association** (et à la quantifier) statistiquement significative entre l'exposition à un facteur et la survenue d'un événement - **Etiologiques :** population en bonne santé, recherche de **facteurs de risque ou protecteurs** de maladie, pour mise en place de mesure de prévention primaire ou secondaire - **Pronostiques :** population de sujets malades, recherche des **facteurs pronostiques** (influençant la survie des malades ou associés à la réponse à un traitement ou à la récidive), pour mise en place de mesure de prévention tertiaire - Exposition mesurée à l'inclusion +/- lors du suivi ; Survenue des événements enregistrée au cours du suivi
	Enquêtes de cohorte		- Prospectives ou rétrospectives ; Maladie et exposition doivent être fréquentes - Sujets indemnes de l'événement étudié à l'entrée dans la cohorte, suivi identique entre exposés et non-exposés
		Enquêtes	**Enquêtes sur échantillon :** sujets extraits par sondage de la population source - Echantillon représentatif : résultats généralisables à la population cible - Echantillon non représentatif : extrapolation des résultats difficile

Enquêtes analytiques

Enquêtes de cohorte

Enquêtes

Enquêtes de type exposés/non-exposés : suivi de 2 cohortes, une de sujets exposés au facteur étudié, l'autre non, avec recherche de survenue de l'événement étudié au cours du temps
- Pas d'estimation directe de l'incidence de la maladie
- Exposition rare et événement fréquent
- Pas d'analyse de plusieurs facteurs de risque simultanément

Statistiques

- Comparaison de **l'incidence** de l'événement entre exposés et non-exposés
- Tenir compte du **risque de base** = risque chez les non-exposés
- Risque = **probabilité de survenue d'un événement** dans une population pendant une période donnée, estimé par la fréquence de l'événement dans l'échantillon
- Mesures d'association : **risque relatif, odds ratio**, **hazard ratio** (tient compte du temps de suivi de chaque sujet, s'interprète comme un RR)

Mesures d'association

Risque en excès = excès de risque : l'exposition ajoute un risque
- Différence entre le risque absolu (probabilité d'être malade si on est exposé) et le **risque de base** (probabilité d'être malade si on n'est pas exposé) ; c'est la **réduction absolue du risque (RAR)** dans les essais

Tableau de contingence :

	Malades	**Non malades**	
Exposés	a	b	a + b
Non exposés	c	d	c + d
	a + c	b + d	

RR = [a/a+c] / [b/b+d]
OR = ad/bc

Risque relatif (RR) : l'exposition multiplie le risque
- Rapport de la probabilité d'être malade si on est exposé sur la probabilité d'être malade si on n'est pas exposé
- Maladie significativement plus fréquente chez les exposés = **association positive = facteur de risque** ; dans le cas inverse => **association négative = facteur protecteur**

Autres expressions possibles :
- **Augmentation relative du risque :** RR-1 => Si le RR est > 1
- **Réduction/Augmentation absolue de risque (RAR) = Différence de risque (DR) : p0-p1 =** différence entre la probabilité d'être malade si on est exposé (p1) et la probabilité d'être malade si on n'est pas exposé (p0). Prend sa valeur entre -1 et 1.
- **Réduction relative du risque (RRR)** : 1-RR (autre manière d'exprimer le RR)
- **Nombre de sujets à traiter (NST)** : nb moyen de sujets à traiter pour éviter un événement = inverse de la différence des risques

Odds ratio (OR) :
- Calcul possible (suivi court ou date exacte arrivée de l'événement non connue), mais n'est pas la méthodologie de référence

Hazard ratio (HR) : tient compte du temps de suivi des sujets ; s'interprète comme un RR

Coefficient de corrélation linéaire (r) : Prend sa valeur entre -1 et 1.
Est d'autant plus proche de 1 (en valeur absolue) que la relation entre les variables est linéaire (y = a x + b avec a ≠ 0)
Coefficient de détermination (r^2) : Prend ses valeurs entre 0 et 1.
Correspond au carré du coefficient de corrélation, évalue la proportion de variabilité d'une variable expliquée par l'autre

Exemple *Essai contrôlé randomisé comparant le risque d'évènement cardiovasculaire à 5 ans entre un groupe traité par un anticholesterolémiant et un groupe traité par placebo en prévention primaire. Résultats (données fictives)*
- *Taux d'évènement à 5 ans :*
  - *Groupe contrôle (placebo) (p0) : 0,05 (5%)*
  - *Groupe expérimental (p1) : 0,03 (3%)*
- *Différence de risque : p0 – p1 = 0,02 (2%)*
- *Risque relatif : p1/p0 = 0,6*
- *Réduction absolue du risque : 1-RR = 0,4 (40%)*
- *Odds ratio : [p1/(1-p1)]/ [p0/(1-p0)] = 0,59*

Nombre de sujets à traiter pour éviter un évènement cardiovasculaire : 1/(p0-p1) = 50

<table>
<tr><td rowspan="4">Enquêtes analytiques</td><td rowspan="3">Enquêtes cas-témoins</td><td colspan="2">- Rechercher dans le passé de deux groupes de sujets, les cas et les témoins, l'exposition antérieure à un ou plusieurs facteurs
- Comparer la prévalence des expositions entre les deux groupes
- Pas de suivi : ce ne sont pas des études longitudinales
- Recueil de l'information sur l'exposition toujours rétrospectif</td></tr>
<tr><td>Sélection des sujets</td><td>- Sur le fait d'avoir ou non l'événement étudié, en insu des expositions
- Effectifs des cas et témoins fixés a priori : estimation impossible de la prévalence de la maladie et de l'exposition
- Appariement des cas aux témoins : rendre les groupes « similaires » et comparables
<u>Cas</u> : malades
- Représentatifs de la population des malades
- Cas incidents, surviennent après la mise en place de l'enquête ; Cas prévalents, sont présents au moment de la mise en place de l'enquête
<u>Témoins :</u> non malades
- Représentatifs de la population d'où sont issus les cas</td></tr>
<tr><td>Statistiques</td><td>- Pas de calcul possible du RR car témoins et cas fixés a priori
- <u>Odss ratio</u> = estimation du RR = rapport des chances ou rapport de cotes d'être malade ou non selon l'existence ou l'absence d'exposition. C'est un rapport de ratios
- Association positive = facteur de risque si exposition plus fréquente chez les cas que les témoins
- Association négative = facteur protecteur, si exposition moins fréquente chez les ca que chez les témoins
- OR surestime le RR, mais bonne estimation si maladie rare ou force d'association faible</td></tr>
<tr><td>Enquêtes transversales</td><td colspan="2">- Principe identique aux études transversales descriptives, mais séquence temporelle impossible à déterminer</td></tr>
<tr><td rowspan="2">Enquêtes expérimentales</td><td>Enquêtes expérimentales</td><td colspan="2">- Enquête où l'attribution (randomisée ou non) de l'exposition (un médicament, une stratégie médicale) est contrôlée par l'investigateur
- Si randomisées = plus haut niveau de preuve scientifique pour montrer l'association causale
- Réalisables que si hypothèse testée est un effet bénéfique de l'intervention (raisons éthiques)
- Statistiques : cf. études de cohorte +/- comparaison de moyennes</td></tr>
<tr><td>Enquêtes quasi-expérimentales</td><td colspan="2">- Pour évaluer des programmes de santé publique ou des pratiques médicales, quand l'allocation de l'intervention ne peut pas être randomisée
<u>Types d'études :</u>
- Avant/après : fréquence événement étudié comparée dans un groupe de sujets avant et après la mise en place de l'intervention
- Ici/ailleurs : fréquence événement étudié comparée entre des sites où les sujets reçoivent l'intervention (ici) et des sites où cette intervention n'est pas pratiquée (ailleurs)
- Avant/après avec groupe témoin : l'évolution de la fréquence de l'événement étudié avant et après intervention est comparée à celle d'un groupe témoin sur la même période</td></tr>
<tr><td rowspan="3">Études diagnostiques</td><td colspan="3">- Etude de la validité d'un nouvel examen diagnostique par rapport à la procédure diagnostique de référence (= Gold Standard) ; Le Gold standard permet de classer les sujets en malades ou non-malades. Tous les sujets bénéficient du Gold standard et du nouveau test à l'étude, idéalement en aveugle l'un de l'autre
- Etudes transversales ++ sur une population de sujets recrutés de façon consécutive, incluant tout le panel des malades possibles ; parfois études cas-témoins mais moindre qualité (pas de calcul de VPP/VPN)</td></tr>
<tr><td>Statistiques</td><td colspan="2">- 4 groupes de sujets : vrais positifs (VP), faux positifs (FP), vrais négatifs (VN), faux négatifs (FN)
<u>Tableau de contingence</u> :
<table>
<tr><th></th><th colspan="2">-- Gold standard --</th></tr>
<tr><th>-- Test à l'étude --</th><th>Malades</th><th>Non-malades</th></tr>
<tr><td>Positif</td><td>VP</td><td>FP</td></tr>
<tr><td>Négatif</td><td>FN</td><td>VN</td></tr>
</table></td></tr>
<tr><td>Propriétés intrinsèques</td><td colspan="2">- Indépendantes de la prévalence de la maladie dans la population
- Sensibilité : probabilité que le nouvel examen soit positif si l'on est malade -> VP/(VP+FN)
- Spécificité : probabilité que le nouvel examen soit négatif si on n'est pas malade -> VN/(FP+VN)
- Rapport de vraisemblance positif (RVP) : rapport de la probabilité que le test soit positif chez un patient malade sur la probabilité que le test soit positif chez un patient sain
- Rapport de vraisemblance négatif (RVN) : rapport de la probabilité que le test soit négatif chez un patient malade sur la probabilité que le test soit négatif chez un patient sain</td></tr>
</table>

Études diagnostiques	Propriétés extrinsèques	- **Dépendent de la prévalence** de la maladie dans la population - **Valeur prédictive positive (VPP)** : probabilité d'être malade si le nouvel examen est positif → VP/(VP+FP) - **Valeur prédictive négative (VPN)** : probabilité de ne pas être malade si le nouvel examen est négatif → VN/(VN+FN)
	Evaluation apport décisionnel	- **Probabilité pré-test qu'un individu soit malade = a priori** - **Probabilité post-test qu'un individu soit malade = révisée a posteriori en fonction des résultats du test** - **Nomogramme de Fagan (= de Bayès) : représentation graphique de la probabilité post-test d'être malade fonction de la probabilité pré-test, du résultat du test et de rapports de vraisemblance**
Schémas moins classiques	**Enquête cas-témoins, nichée dans une cohorte :** - Cas et témoins sélectionnés parmi les sujets d'une cohorte déjà suivie - Respect de la séquence chronologique, information de meilleure qualité car recueil prospectif des données **Enquête cas-croisé :** - Enquête cas-témoins où chaque cas est son propre témoin ; Enquête observationnelle, différente des essais randomisés en *cross-over* - Exposition d'un sujet dans un délai proche de l'événement étudié (**période cas**) est comparée à celle d'une période éloignée de la survenue de l'événement où ce même sujet ne présentait pas l'événement (**période témoin**) - Permet d'analyser les effets d'une exposition ponctuelle/intermittente sur le risque de survenue d'un événement aigu - Limite les **facteurs de confusion** et **les biais** liés à la sélection des témoins **Enquête cas-cohorte :** - Présente les avantages des études de cohorte et cas-témoins, avec possibilité d'étudier plusieurs événements d'intérêt à moindre coût que les cohortes classiques - Sélection d'un échantillon représentatif de la cohorte (= « sous-cohorte »), indépendamment de l'apparition de la maladie, avec recueil dans le temps des **cas incidents** de la cohorte et de la sous-cohorte - Comparaison des cas incidents (cohorte + sous-cohorte) avec les non-cas de la sous-cohorte **Etude sur données groupées = agrégées :** - Estimation de la fréquence du phénomène de santé au niveau **d'un groupe** et non de manière individuelle - Comparaison de ces fréquences entre plusieurs populations ou dans le temps	

INTERPRÉTATION

Conclusion erronée des résultats pouvant être due à des :
- **Erreurs aléatoires** dues aux **fluctuations d'échantillonnage** (hasard) ou à un **manque de puissance** ; ces erreurs sont diminuées si la taille de la population augmente et sont prise en compte dans la présentation des résultats (IC95%, p)
- **Erreurs systématiques** ou **biais** (non dues au hasard), se surajoutent aux erreurs dues au hasard

Interprétation des résultats en trois étapes : **jugement de signification statistique**, **estimation des risques**, **biais** :
- En cas d'association significative : discuter les erreurs aléatoires et les biais
- En cas d'association non significative : discuter un manque de puissance et les biais

Jugement de signification statistique

= erreur aléatoire dans une **étude explicative**
- Recherche l'existence d'association entre une/des exposition(s) et un/des événement(s) par **comparaison des groupes** avec des tests appropriés (chi^2, Student)

Résultat du test **significatif** si la valeur de ***p*** (« ***p-value*** », **degré de signification**) est inférieure à un seuil déterminé *a priori* (**risque alpha**, en général 5%). Le risque alpha est le risque, fixé *a priori*, de conclure à tort à une différence, alors que la valeur de *p* mesure ce même risque *a posteriori*
- Degré de signification reflète **la confiance** (matérialisée par **l'intervalle de confiance**) accordée aux résultats mais insuffisant à leur interprétation (les facteurs de confusion potentiels et les risques d'erreurs doivent avoir été contrôlés) ; il informe sur la **force de l'association** et **la précision** de la mesure

Résultat du test **non significatif** : s'interroger sur **la puissance**
- La puissance dépend de **la taille de l'échantillon**, de **la différence entre les groupes**, **du risque alpha** (plus ces paramètres sont élevés, plus la puissance augmente), de la **variabilité des observations** (si la variance augmente, la puissance diminue), et selon **l'uni ou bilatéralité du test** (unilatéral = puissance augmentée)
- Puissance ***a priori*** intervient dans le calcul du NSN (avec d'autres hypothèses) ; Puissance ***a posteriori*** discutée en cas de résultats négatifs
- Manque de puissance = probabilité de conclure à tort à une différence non significative alors qu'il existe une association dans la « vraie vie » = erreur de **second espèce** ou **risque bêta**

Biais

= **Erreur systématique** dans l'estimation d'un paramètre, pouvant masquer, renforcer voire créer une liaison entre un facteur et une maladie → **sur- ou sous-estimation du risque**
- Des biais majeurs peuvent invalider totalement des résultats
- Les biais ne peuvent être diminués en modifiant la taille de l'échantillon

Biais de sélection

Lié au mode de **sélection** des sujets (constitution de l'échantillon) ou du **suivi** des sujets (**biais d'attrition**)
Responsable d'une **absence de représentativité** de la population cible (*recrutement hospitalier, critères d'inclusion trop sélectifs, défaut de participation, données manquantes, etc.*)
Non différentiel = problème de validité externe
Différentiel = sélection sujets différente selon groupe

Conséquences
- Enquêtes transversales : estimations de prévalence faussées
- Enquêtes comparatives : distorsion des estimations de risque du fait d'une comparabilité des groupes altérée

Prévention
- Au moment de la conception car ne peuvent être contrôlés *a posteriori*

Etudes cas-témoins :
- Témoins doivent être représentatifs des non-malades de la population source d'où sont issus les cas
- Sélection des cas et témoins **en aveugle** de l'exposition

Cohorte :
- Sujets non exposés doivent être représentatifs des sujets non exposés de la population source
- Minimisation des perdus de vue + en faire une analyse qualitative et quantitative

Biais de classement

- Lié à la mesure de l'information : **biais de classement, de mesure ou d'information**
- Problème de mesure de l'exposition (cas-témoins) ou dans la détermination de la maladie (enquête de cohorte), responsable d'un mauvais classement des sujets et donc d'une mauvaise estimation des indicateurs de santé ou des mesures d'association

Biais de classement non-différentiel :
- Erreur de mesure ne diffère pas entre les groupes (ex. : prb étalonnement appareil de mesure)
- **Sous-estimation** de la force de l'association entre E et M

Biais de classement différentiel :
- Erreur de mesure diffère entre les groupes (information sur exposition recueilli en connaissance du statut cas ou témoins, recherche plus régulière de la maladie chez exposés que non-exposés
- **Sous-estimation** ou **surestimation** (le plus souvent) du risque

Prévention
- Au moment de la conception car ne peuvent être contrôlés *a posteriori*, au moment de l'analyse
- Cas-témoins : interrogatoire standardisé pour recueil des expositions, enquêteurs en aveugle des hypothèses testées
- Cohorte : investigations strictement similaires entre sujets exposés et non exposés

Biais de confusion

- Lié à la présence de **facteurs de confusion** (liés à la fois à l'événement et au facteur de risque étudié), conduisant à une **estimation erronée (surestimation ou sous-estimation) de la mesure d'association** entre les deux
- N'existe pas dans les études descriptives et les études diagnostiques : **uniquement pour analytiques**
- Possible « confusion résiduelle » dans les études observationnelles du fait de la non prise en compte de l'ensemble des facteurs de confusion potentiels
- Les techniques de prises en compte de ce biais (cf. ci-dessous) peuvent être combinées dans une même étude : **planifiés** au moment de la conception de l'étude et/ou **corrigés** lors de l'analyse

Randomisation
- Technique **la plus efficace**
- Permet de rendre comparables les groupes à l'inclusion sur tous les facteurs de confusion possibles, **connus et inconnus**

Appariement
= Ajustement *a priori* (cas-témoins et exposés/non-exposés)
- Au moment du recrutement, pour rendre comparables les sujets sur les variables d'appariement
- Appariement strict (ex : sexe), ou non-strict (ex : âge ± 5 ans)

Ajustement
- Lorsque plusieurs facteurs sont associés à l'événement étudié
- Teste l'association **après ajustement** sur les facteurs de confusion potentiels : **analyses multivariées** par **régression logistique** (OR), **régression linéaire** (moyennes) ou **modèle de Cox** (HR)
- On ne prend en compte que les facteurs sur lesquels on a ajusté
- Permet d'identifier l'impact des facteurs d'exposition étudiés, toute chose étant égale par ailleurs = identification des facteurs **indépendamment** liés au phénomène étudié
- Mesures d'association non ajustées = **brutes**

<table>
<tr><td>Interaction</td><td colspan="3">- Permet de déterminer si le risque d'événement lié à un facteur diffère en fonction du niveau d'exposition à un autre facteur
- On parle d'interaction lorsque l'association entre une exposition (E) et une maladie (M) varie en fonction du niveau d'un autre facteur (F). L'OR ou le RR n'a pas de sens. Nécessité d'une analyse stratifiée
- On parle de confusion si la liaison de F avec E et M peut entraîner un biais de confusion dans la liaison entre E et M ; Nécessité d'une analyse multivariée</td></tr>
<tr><td rowspan="3">Causalité</td><td colspan="3">- Causalité discutée si l'exposition est associée à une augmentation significative du risque de maladie et que l'étude ne présente pas de biais majeurs
- Dans une enquête observationnelle, la causalité est d'autant plus probable que plusieurs des critères de Bradford-Hill sont réunis (faisceau d'arguments, tous les critères étant rarement réunis)</td></tr>
<tr><td rowspan="2">Critères de Bradford-Hill</td><td>Internes à l'étude</td><td>- Temporalité : l'exposition précède la maladie
- Force de l'association : plus la force de l'association est grande, moins l'association est susceptible d'être expliquée par des facteurs de confusion
- Spécificité de l'association : l'exposition étudiée est présente chez tous les malades (ou presque) et seulement chez eux, et pour cette seule maladie
- Relation dose-effet : plus l'exposition augmente, plus on a de risque de développer la maladie
- Cohérence interne : schéma (méthodologie) et prise en compte des biais</td></tr>
<tr><td>Externes à l'étude</td><td>- Constance de l'association : et reproductibilité dans différentes situations (différentes périodes, régions ou populations)
- Plausibilité biologique : existence d'un mécanisme d'action de l'exposition sur la maladie
- Cohérence externe : reproductibilité des résultats
- Preuves expérimentales : une intervention visant à diminuer le facteur de risque doit faire diminuer l'incidence de la maladie
- Analogie : raisonnement par analogie = une autre association entre une exposition analogique et une maladie analogique</td></tr>
<tr><td rowspan="4">Impacts en population</td><td colspan="3">- Un RR ou un OR significatif ne reflète pas forcément que l'exposition ait un impact très important en population générale, car cela dépend de la fréquence de l'exposition (= mesure d'impact) et de la taille de l'effet
- Facteur d'exposition = facteur d'intérêt dans une étude expérimentale ou analytique :
- Etude analytique = facteur de risque (FDR) = associé à une probabilité plus élevée de survenue du phénomène de santé (définition statistique et non causale)
- Etude étiologique = facteur étiologique = nature causale du FDR sous-entendue/supposée/prouvée
- Etude pronostique = facteur pronostique
- Déterminants de santé influencent l'état de santé d'une population sans être des causes directes de maladies</td></tr>
<tr><td rowspan="2">Fréquence de l'exposition</td><td>Risque attribuable</td><td>- Proportion de cas attribuables au facteur :
RA = [F x (RR-1)] / [F x (RR-1) +1]
avec F fréquence d'exposition
- Le risque attribuable n'a de sens que si la relation causale est probable
- Dans les études cas-témoins, RR estimé par OR, et F = prévalence de l'exposition chez les témoins
- RA x prévalence = estimation du nombre de cas qui seraient évités si l'exposition était supprimée</td></tr>
<tr><td>Fraction étiologique</td><td>- Proportion de cas attribuables au facteur parmi les exposés
= fraction de risque attribuable chez les exposés
FE = (RR-1) / RR</td></tr>
<tr><td>Taille de l'effet</td><td colspan="2">- Enquêtes réalisées avec des groupes de taille de plus en plus importantes, pour lesquels de petites différences entre les groupes peuvent être significatives
- Raisonner en termes de taille d'effets et de pertinence clinique</td></tr>
</table>

Littérature scientifique

Publication	- Résultats valorisés et partagés avec la communauté scientifique (communications en Congrès, articles scientifiques - Processus de publication avec **comité de lecture** - Article soumis à des pairs – experts du domaine – pour évaluation de la qualité (« **peer-review** ») : article refusé, accepté en l'état, accepté avec modifications - Attention aux **revues prédatrices** (escroquerie sur publications moyennant paiement)	
Niveau de preuve	**GRADE DE RECOMMANDATION**	**NIVEAU DE PREUVE SCIENTIFIQUE**
	A) Preuve scientifique établie	**Niveau 1** . Essais comparatifs randomisés de forte puissance . Méta-analyse d'essais comparatifs randomisés . Analyse de décision fondée sur des études bien menées
	B) Présomption scientifique	**Niveau 2** . Essais comparatifs randomisés de faible puissance . Etudes comparatives non randomisées bien menées . Etudes de cohortes
	C) Faible niveau de preuve scientifique	**Niveau 3** . Etudes cas-témoins.
		Niveau 4 . Etudes comparatives comportant des biais importants . Etudes rétrospectives . Séries de cas . Etudes épidémiologiques descriptives (transversale, longitudinale)